"Tensor Regression with Applications in Neuroimaging Data Analysis"

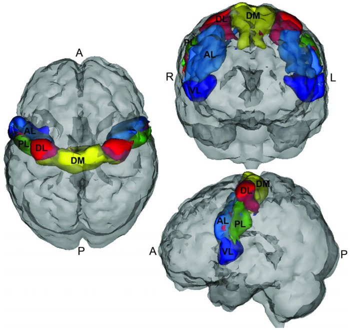
Hua Zhou, Lexin Li, & Hongtu Zhu

Jennifer Kirk

April 17, 2013

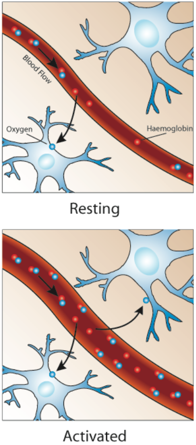
Scientific Motivation

- Mental health disorders are difficult to diagnose and treat
- Physiology of the brain is not well understood
- Neuroimaging can elucidate the brain's physiology
- Several types of neuroimaging, e.g. PET, MRI, **fMRI**

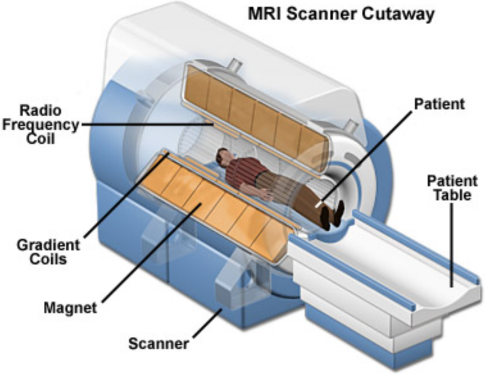


Brain Areas Associated with ADHD from fMRI
Image Source: MIT Tech Review

BOLD & fMRI



Blood Oxygen Level Dependent Effect (BOLD)
Image Source: FMRIB Oxford



MRI Scanning Machine
Image Source: Discover Magazine

fMRI Data

- Slice: Section of the brain in 1-D
- Voxel: 3-D pixel, $\approx 3\text{mm}^3$
- Measure BOLD signal at each voxel at each time point
- Final data is a 4-D array

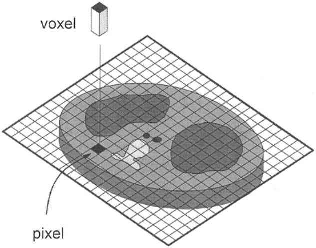
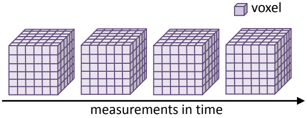


Image Source: Victor J. Montemayor



Statistical Motivation

- Want to use **fMRI images as a covariate**
- Naive approach: use image as vector covariate
  - Lots of data $\Longrightarrow$ **lots of parameters** ($\approx$ 16 million)
  - Ignores **spatial and temporal correlation**
- This paper tries to solve this problem
- Current Methods
  - Voxel-Based Methods
  - Functional Data Methods
  - Two-Stage Reduction Methods

Current Methods

- Voxel Based Methods
  - Analysis of each voxel as response variable
  - Assumes voxels independent–ignores spatial correlation
- Functional Data Methods
  - Collapses data into one parameter function
  - Commonly used for 2-D data, extension to 3-D data is complex
- Two-Stage Reduction Methods
  - Reduce the dimension of the data, possibly more than once, then model the reduced data
  - Theoretical properties are intractable and reduction maybe unrelated to response

Rank-R Generalized Linear Tensor Regression

Extend the GLM to tensor covariates by assuming you can decompose the tensor parameter into a simple form and estimate

GLM framework (McCullagh & Nelder)

- Outcome $Y \sim$ exponential family
- Vector Covariate: $\mathbf{Z}$
- Link function: $g(\mu) = \alpha + \gamma^\mathsf{T}\mathbf{Z}$

Rank-R Generalized Linear Tensor Regression

Extend the GLM to tensor covariates by assuming you can decompose the tensor parameter into a simple form and estimate

Rank-R Generalized Linear Tensor Regression

- Outcome $Y \sim$ exponential family
- Vector covariate: $\mathbf{Z}$, Tensor covariate: $\mathbf{X}$
- Matrix Covariate (D $=$ 2)
  - Link function: $g(\mu) = \alpha + \gamma^{\mathsf{T}}\mathbf{Z} + \beta_1^{\mathsf{T}}\mathbf{X}\beta_2$
- Tensor Covariate (D $>$ 2)
  - Assume $\mathbf{B} = [[\mathbf{B}_1, \ldots, \mathbf{B}_D]] = \sum_{r=1}^{R} \beta_1^{(r)} \circ \ldots \circ \beta_D^{(r)}$
  - Link function: $g(\mu) = \alpha + \gamma^{\mathsf{T}}\mathbf{Z} + \langle(\mathbf{B}_D \odot \ldots \odot \mathbf{B}_1)\mathbf{1}_R, \mathrm{vec}(\mathbf{X})\rangle$

Rank-R Generalized Linear Tensor Regression

- Maximum Likelihood Estimation
- Estimation Algorithm
  1. Estimate the vector covariate parameters, assuming $\mathbf{B} = 0$
  2. Estimate each tensor covariate parameter Bd, using current estimates for $\mathbf{B}_k$, $k < d$ and previous estimates for $\mathbf{B}_k$, $k > d$
  3. Iterate until the likelihood converges
- Authors show identiability, consistency and asymptotic normality

Summary

- Analysis of complex neuroimages is important for understanding brain physiology
- fMRI data is complex: 4-D array with spatial and temporal correlation
- Current analysis methods ignore one or more of these features
- Tensor regression, an extension of GLM to array covariates allows efficient analysis without ignoring these features
- Next Steps: literature review of current methods, simulations & proofs